

Section 4.2.2 Los Alamos Computer Science Institute

Project Leads:

Andy White, LANL, (505) 665-4700, abw@lanl.gov

Deputy Project Lead:

John Thorp, LANL, (505) 665-2000, thorp@lanl.gov

Purpose and Description: The Los Alamos Computer Science Institute (LACSI) has the overarching goal of fostering computer science research efforts that are both internationally recognized and relevant to the goals of the Los Alamos National Laboratory (Los Alamos). LACSI was formed to:

- Build a presence in computer science research at Los Alamos that is commensurate with the strength of the physics community at Los Alamos,
- Achieve a level of prestige in the computer science community that is on a par with top computer science departments in the nation,
- Pursue computer science research that is relevant to the goals of the ASCI program at Los Alamos, and
- Ensure that there remains a strong focus on high-performance computing in the academic computer science community.

LACSI will achieve these objectives through a combination of strategic and tactical goals engaging top laboratory and university computer scientists. In particular, a majority of the research effort in LACSI will focus upon three strategic, long-term research efforts of critical importance to the future of the ASCI program: (1) Application Reliability, (2) Application Performance, and (3) Application Development. As progress is made against these research objectives, the long-term needs of the ASCI program will be assessed and the strategic goals of the Institute re-evaluated. LACSI firmly believes that to make a difference on the research horizon, a large, focused, long-term effort must be brought to bear on a research discipline critical to the success of the ASCI program—LACSI is exactly the vehicle to enable this quality of research effort.

Finally, for the Institute to be successful, it must build strong academic ties to experts in the community able to further the strategic goals of the Institute in support of the ASCI program. Our collaborations include Rice University, University of Houston, University of Tennessee, University of New Mexico, and University of Illinois–Urbana Champaign. Detailed mechanisms of interaction and management are described below.

Collaboration, Coordination, and Management: LACSI will have a physical instantiation at Los Alamos, rather than a virtual existence across multiple organizations. To enable this objective, LACSI will find its home in the Los Alamos Advanced Computing Laboratory (ACL). The ACL is already a recognized leader in many computer science disciplines and is an active

member in the Office of Science, National Science Foundation, and Defense Program computer and computational science research communities. The LACSI presence in the ACL will stretch research horizons to the long-term objectives necessary to enable 100-teraOPS simulation in a five-year timeframe. Likewise, LACSI will benefit from building upon an organization that already has: a strong computer science presence at the Laboratory; a strong presence in the external community; a healthy student and visitors program; sponsored seminar series and symposia; a physical building in the open to collocate efforts; and an administrative staff to coordinate activities and publicize events.

Some Institute efforts are joint between the universities and Los Alamos, other efforts are focused at universities, and other efforts at Los Alamos. One of the principal objectives of the Institute is to leverage the strengths of Los Alamos and its university partners to create a world-class organization focused upon the strategic research objectives of the Institute. LACSI does not seek to mirror all activities between Los Alamos and its university partners—each participant in the Institute will focus upon its core research strengths and engage other Institute participants in multidisciplinary collaboration. To ensure the necessary interaction between Los Alamos and university researchers and to build upon their strengths, LACSI will use the LACSI Collaboration investment in addition to the travel budgets accorded Los Alamos and university researchers to ensure the necessary interaction through an active visitor, seminar, and conference program.

The Institute will be managed by a Director, Andy White, who will be guided by an executive committee that consists of Andy White, (Co-Chair); Ken Kennedy, Rice University (Co-Chair); Dan Reed, University of Illinois at Urbana-Champaign (UIUC); Rick Stevens, Argonne National Laboratory; Lennart Johnson, University of Houston; Jack Dongarra, University of Tennessee; Dan Sorensen, John Mellor-Crummey, Rice University; John Thorp, Los Alamos, and Linda Torczon, Rice University. The executive committee will convene quarterly to track progress against research objectives. Furthermore, the executive committee will convene an annual review to evaluate the research efforts from the point of view of computer science.

LACSI will integrate efforts with the other ASCI-supported institute efforts. In particular, Los Alamos and Lawrence Livermore National Laboratory (Lawrence Livermore) will coordinate solver and run-time systems research to avoid duplication of effort and build upon the strengths of each Institute. In particular, the Center for Advanced Scientific Computing (CASC) will invest in solver research as a core competency and LACSI will invest in scalable run-time systems as a core competency. In turn, Los Alamos will invest to a smaller degree in solver research with the explicit goal of collaborating with CASC and assimilating research results into ASCI simulation efforts. Likewise, CASC will invest in run-time systems research with the similar goal of inserting LACSI run-time system research into Lawrence Livermore simulation efforts. This collaborative approach enables the ASCI program to build upon the diverse strengths of research efforts at the National Laboratories while avoiding duplication of effort.

University Contracts: LACSI supports external research at a number of universities through a contract with HiPerSoft at Rice University. This contract is a major portion of the Institute's budget and correspondingly contributes to several of LACSI's research sub-projects.

Fellowships are also handled through a contract relationship with the Krell Institute.

Issues, Constraints, and Assumptions: One of the principal constraints is the ability to attract and retain talented computer scientists. LACSI's ability to retain talented scientists is looming as a large problem.

Another assumption is continued support from the Office of Science through the restructuring of defense program activities into a semi-autonomous agency. Many of the research efforts in the Institute are also supported by Office of Science investment. Loss of this support would affect the milestone schedule described below. The Laboratory Directed Research and Development (LDRD) investments are long-term investments and the problem-solving environment (PSE) program currently views the aligned contributions into LACSI areas of research as critical investments to guarantee an insertion path of technologies into ASCI simulations efforts. Nonetheless, a sudden significant reduction in PSE and LDRD investment would also impact the schedule of tasks and milestones described below.

4.2.2.1 Application Performance

Project Lead:

Adolfy Hoisie, LANL/CCS-3, (505) 667-5216, hoisie@lanl.gov

Sub-Project Description: Extreme-scale parallel architectures require novel approaches to achieving high performance on applications. The increased complexity of the machines, the depth of their memory hierarchies, the heterogeneity of the interconnect, and the sheer level of parallelism that the hardware and system software exposes, pose significant challenges to achieving high application performance and scalability. This sub-project will address in a coherent fashion some of the important issues related to achieving application performance. To this end, section 4.2.2.1.1 "Performance Tools" is concerned with dynamic instrumentation of applications and building a performance validation infrastructure. The research in 4.2.2.1.2 "Performance Analysis and Modeling," will continue the performance analysis and modeling of the ASCI workload on the ASCI platforms. In 4.2.2.1.3 the emphasis will be on the development of Java-based tools for rapid prototyping of algorithms. The important issue of single-node performance will be addressed in 4.2.2.1.4 within the context of compiler and run-time assisted tools for code optimization.

The work in the area of "Application Performance" will be conducted using an ASCI-characteristic workload and will target architectures and software infrastructure of interest to ASCI. Most of the work will involve continuous collaboration with the ASCI application teams at Los Alamos. The findings from this research, the tools and the software infrastructure are likely to benefit all ASCI application teams through improved performance, better performance instrumentation capabilities, and improved algorithm-architecture mapping.

4.2.2.1.1 Performance Tools

Sub-Project Leads:

Dan Reed, University of Illinois at Urbana-Champaign, (217) 333-3807, reed@cs.uiuc.edu

Jeff Brown, Los Alamos National Laboratory, jeffb@lanl.gov

Sub-Project Description: The challenge of achieving high performance is exacerbated by the need to integrate code and libraries from multiple developers when attacking complex, realistic problems (e.g., integrating structures, materials, and fluid dynamics).

To increase performance scalability, continued developing a suite of performance contract software at UIUC is proposed, targeting large-scale clusters. These performance contracts rely on characterization of application performance signatures (application stimuli to resources), resource behavioral profiles (expected resource responses to application stimuli), and a runtime measurement and performance validation infrastructure that compares expected and observed behavior. This work will focus on clusters of the type being deployed by ASCI and related projects (e.g., the NSF PACI program) and the interplay of libraries and multidisciplinary application code. It will build on the SvPablo and Autopilot toolkits.

Los Alamos efforts will focus on the development of critical performance tool capabilities that provide insight into barriers to application code scalability. Tool development will leverage a portable tool infrastructure developed in collaboration with the proposed Maryland Level 2 alliance. The Los Alamos tool component will be built on the infrastructure to provide for dynamic run-time extraction and analysis of performance data. A “performance debugger” will be developed that will allow the user precise control over what performance data is extracted and analyzed. For example, the tool will allow the user to run a parallel job to a specific point of interest (say after problem setup), insert probes, run a few computational cycles to extract performance data, analyze the data, then decide what other data is of interest to drill into the reasons behind scalability barriers. This proposed tool capability will leverage the knowledge a user has of application dynamics and make large-scale performance analysis a trackable problem due to techniques to limit the amount of data be analyzed.

Collaboration and Coordination: With regard to the tasks discussed below, no indication is given as to the ASCI mileposts supported, since the association of certain ITS milestones with ASCI mileposts presents possible classification issues. In all cases, however, ITS research and development activities are driven by ASCI code needs.

University Contracts: This work will be performed in a university setting at UIUC.

Milestones and Associated Tasks:

Tasks:

- UIUC
 - SvPablo extensions to display contract validation (Quarter 3).

- Contract extensions for I/O, communication, and computation within multilevel library and discipline context (Quarter 4).
- Validation with application codes (throughout the year).
- Los Alamos
 - Initial release of the “Los Alamos Performance Debugger” (Quarter 4).
 - Initial deployment of the tool infrastructure (Quarter 2).

4.2.2.1.2 Performance Analysis and Modeling

Sub-Project Lead:

Adolfy Hoisie, LANL/CCS-3, (505) 667-5216, hoisie@lanl.gov

Sub-Project Description: The task of performance analysis and modeling of the ASCI workload will continue. Building on the novel methodology developed at Los Alamos, this project will expand the scope of the work in the challenging direction of modeling adaptive, stochastic, and/or nondeterministic computations. A key result from this work should be a detailed *a-priori* understanding of how load-balancing techniques affect runtime.

In conjunction with system architecture research, in the architectures (communication subsystem, system integration, I/O, etc) of interest to ASCI are analyzed and modeled, this work results in predictive performance and scalability models of our ASCI workload. Such results are essential to a complete understanding of application/architecture interaction, and should also indicate ways for achieving higher application performance and better system utilization.

Collaboration and Coordination: With regard to the milestones discussed below, no indication is given as to the ASCI mileposts supported, since the association of certain ITS milestones with ASCI mileposts presents possible classification issues. In all cases, however, ITS research and development activities are driven by ASCI code needs.

University Contracts: Collaborations with UIUC (Professors Torrellas and Reed), UCLA (Professor Bagrodia), Texas A&M (Professor Raushwerger), Rice University (Professor John Mellor-Crumey) and University of Houston (Professor Barbara Chapman).

Milestones and Associated Tasks:

Tasks:

- Inclusion of adaptation in the model of Sage/Rage: Analysis of performance improvement for various possible data structures for computations with Sage using adaptive mesh refinement. (Quarter 2)
- Performance analysis and modeling of MCNP: MCNP is an important application in the ASCI workload. It utilizes a mixed-programming paradigm threads-MPI. Due to the lack of domain decomposition, the memory requirements are very high. Performance of the combination threads-MPI will be a major component of this research. Creation of a

runtime model of MCNP based on which performance prediction of the code on various architecture will be made. Performance with and without domain decomposition will be analyzed. (Quarter 4)

- Performance of ASCI applications on a cluster of Intel IA-64 processors: We will analyze the performance of the ASCI workload on a cluster of IA-64 processors using various high-performance interconnect strategies. Single-node performance on IA-64 processors, as well as the suitability of Itaniums to achieving high-performance I/O will be emphasized. (throughout the year)

4.2.2.1.3 Rapid Performance-Prototyping of Algorithms

Sub-Project Leads:

W. B. VanderHeyden, LANL, (505)667-9099, wbv@lanl.gov

Ken Kennedy, Rice University, (713) 348-5186, ken@rice.edu

Sub-Project Description: The goal of this sub-project is to develop tools for the rapid prototyping of algorithms and models for use in complex large-scale scientific simulation codes. The approach will be to capitalize on two successful efforts—one at Los Alamos and one at Rice University—involving the use of the Java programming language as a basis for rapid prototyping tools. Java has a number of attractive features for the development of rapid prototyping tools, including full support for software objects, parallel and networking operations, relative language simplicity, type-safety, portability, and a robust commercial marketplace presence leading to a wealth of programmer productivity tools.

The Los Alamos effort concerns the development of a pure-Java rapid prototyping environment called CartaBlanca. The thrust of CartaBlanca is to explore synergies among object-oriented programming, flexible Jacobian-Free Newton-Krylov solution methods, Java, and modern software engineering tools, and to use these synergies to produce a rapid method and model prototyping tool for the simulation of coupled nonlinear physics problems on unstructured grids of interest to ASCI.

The effort at Rice University will be coordinated with Cartablanca to develop improved compilers for Java. To enhance performance of object-oriented Java programs, Rice will develop a whole-program transformation tool called JaMake that performs inter-procedural optimizations on Java. This will include two novel whole-program optimizations, “class specialization” and “object inlining,” which can improve the performance of high-level, object-oriented, scientific Java programs by up to two orders of magnitude.

Coordination with Campaigns/Directed Stockpile Work (DSW): With regard to the milestones discussed below, no indication is given as to the ASCI mileposts supported, since the association of certain ITS milestones with ASCI mileposts presents possible classification issues. In all cases, however, ITS research and development activities are driven by ASCI code needs.

Collaboration and Coordination: Collaborations external to Los Alamos will significantly enhance the products from this sub-project. Rice University, through Professor Kennedy's group, will contribute their expertise in rapid prototyping environments, Java compilers, Java code optimization and Java distributed shared memory systems. Interactions with Compaq research will continue to help optimize CartaBlanca's efficiency through improved use of the Java compilers and VM's on the new Los Alamos 30T ASCI computer system.

Issues, Constraints, and Assumptions: Assumptions: (1) development time will be made available on the ASCI Q and Blue Mountain machines for this work, (2) timely assistance from ASCI Q and Blue Mountain system administrators will be available for reasonable requests for Java resources on those machines, (3) a copy of JaMake will be obtained from Rice University in time for testing before the end of the FY, (4) freely available DSMs will be compatible with the CartaBlanca software and our ASCI Q and Blue Mountain operating systems.

Milestones and Associated Tasks:

Tasks:

- Fully implement thread-based, parallel version of CartaBlanca using Java's built-in thread facilities. Ensure load balance for efficient speed-up (Los Alamos) (Quarter 2).
- Incorporate 2-level additive Schwarz or other appropriate pre-conditioner to ensure scalable performance (Los Alamos) (Quarter 2).
- Test parallel version of CartaBlanca on appropriate shared memory system using nonlinear multiphase flow physics on an unstructured grid; document parallel performance improvement (Los Alamos) (Quarter 3).
- Test CartaBlanca within a distributed shared memory (DSM) environment for parallel computations on distributed memory systems (Los Alamos/Rice) – Quarter 4.
- Test appropriate compilers for Java high-performance computing, including the JaMake system, if available. (Los Alamos/Rice) (Quarter 4).

4.2.2.1.4 Compiler Technology for High Performance

Sub-Project Lead:

John Mellor-Crummey, Rice University, (713) 348-5179, johnmc@cs.rice.edu

Sub-Project Description: This project performs research on compiler, run-time, and tools technology to help programs achieve a higher fraction of peak performance on current and future generations parallel systems. Research topics include compiler and run-time technology for improving memory hierarchy utilization, compiler technology for exploiting advanced processor

architectures, and compiler technology for optimizing communication on parallel systems. Based on this research, compiler-assisted tools are being developed for pinpointing performance problems in applications. The tools are applied to ASCI applications to both improve the performance of those codes and to identify opportunities and requirements for future work.

Coordination with Campaigns/ Directed Stockpile Work (DSW): With regard to the milestones discussed below, no indication is given as to the ASCI mileposts supported, since the association of certain ITS milestones with ASCI mileposts presents possible classification issues. In all cases, however, ITS research and development activities are driven by ASCI code needs.

Collaboration and Coordination: The project team will coordinate with the problem-solving environment (PSE) deployment team to deploy current and future versions of the HPCView toolkit and source-to-source transformation tools. The project team will interact with application teams to assist in applying these tools and collaborate in identifying additional opportunities for applying compiler and run-time technology to improve application performance.

University Contracts: This work will be performed at Rice University under the Los Alamos Computer Science Institute (LACSI) contract.

Issues, Constraints, and Assumptions: The HPCView toolkit uses SpeedShop on SGI platforms as the source of its hardware performance counter data; an Alpha port is currently underway. Applicability to other platforms depends on having hardware and operating system support for statistical sampling of program events.

Source-to-source transformation tools for high-level program optimization depend on high-quality vendor compilers for generating efficient machine code.

These goals are contingent on maintaining current staffing levels.

Milestones and Associated Tasks:

Tasks:

- Package and test HPCView, scripts, and documentation for deployment on SGI Origin and Tru64 platforms (Quarter 1).
- Develop scripts and data extraction tools to utilize DCPI/ProfileMe performance data in HPCView on Tru64/EV67 platforms. (Quarter 2).
- Adapt SGI Pro64 compiler infrastructure for source-to-source optimization of FORTRAN 90 codes (Quarter 1).
- Port and integrate program analysis mechanisms into the framework (Quarter 2).
- Add transformations for improving memory hierarchy performance (Quarter 3).

- Test tool on benchmarks and other open codes (Quarter 4).

4.2.2.2 Application Development and Software Subsystems

Project Lead:

Rod Oldehoeft, LANL/CCS-1, (505) 665-3663, rro@lanl.gov

Sub-Project Description: Research for new problem-solving environment (PSE) application-specific software infrastructures based on novel compiler technologies and library designs will be conducted. This work will address issues of high performance and error-prone interfacing common in using scripting languages by using a “telescoping language” approach in which application-specific translators are automatically generated from annotated component libraries. A high-level application development environment for an application of interest to Los Alamos will be constructed, based on MATLAB.

In conjunction, annotated libraries of parallel computational methods relevant to DOE/Los Alamos applications will be developed. These components include nonlinear optimization algorithms, Krylov projection methods for linear systems and eigenvalue calculations, reduced-basis methods for dynamical systems, discrete optimization methods for solving nonsmooth, multiscale coupled models, and inverse control problems. This will also contribute to the infrastructure for developing new applications that require coupling of subsystem simulations.

To ensure that software products such as these developed by Los Alamos Computer Science Institute (LACSI) partners are made available efficiently, common practices and procedures will be developed to ensure appropriate software distribution, i.e., via Open-Source Software licenses or via the intellectual-property procedures in place at partner institutions.

PSEs developed with telescoping languages and numerical libraries designed to work with them represent the next generation of component integration frameworks. Such frameworks can support rapid prototyping of ASCI applications at higher efficiency and reliability than earlier approaches. Efficient distribution of software products under conditions favorable to ASCI and Los Alamos will increase the impact of the software.

4.2.2.2.1 Open Source Software

Project Leads:

**Rod Oldehoeft, LANL/CCS-1, (505) 665-3663, rro@lanl.gov
Danny Powell, Rice University, (713) 348-6011, danny@rice.edu**

Sub-Project Description: The Open-Source Software (OSS) development and distribution model is proving to be a useful vehicle for building and sharing software important to high-performance computing in general and ASCI in particular. LACSI partners frequently produce software products of value to the broader computer science community. Making them available efficiently requires establishing a coordinated set of procedures that each participating institution can integrate into its overall intellectual-property management. This project will capitalize on the OSS release mechanisms developed at Los Alamos, improving and customizing them for use by all the LACSI partners.

Collaboration and Coordination: With regard to the milestones discussed below, no indication is given as to the ASCI mileposts supported, since the association of certain ITS milestones with ASCI mileposts presents possible classification issues. In all cases, however, ITS research and development activities are driven by ASCI code needs.

Milestones and Associated Tasks:

Tasks:

- Discuss Los Alamos procedures with researchers and administrators at partner sites.
- Describe and document Los Alamos OSS distribution policies and procedures.
- Develop common practices and procedures that work for all partner sites.

4.2.2.2 PSE Research for Science and Engineering

Sub-Project Lead:

**Ken Kennedy, Rice University, Center for High Performance Software Research,
(713) 348-5186, ken@rice.edu**

Sub-Project Description: The goal of this project is to develop compiler technologies and library designs that will make it possible to automatically construct domain-specific development environments for high-performance applications. This project will develop advanced compiler technology to construct high-level programming systems from domain-specific libraries. Programs would use a high-level scripting language, such as Matlab, to coordinate invocation of library operations. Scripting languages typically treat library operations as black boxes and thus fail to achieve acceptable performance levels for compute-intensive applications. Previously, researchers have improved performance by translating scripts to a conventional programming language and using whole-program analysis and optimization. Unfortunately, this approach leads to long script compilation times and has no provision to exploit the domain knowledge of library developers.

To address these issues we are pursuing a new approach called “telescoping languages,” in which libraries that provide component operations accessible from scripts are extensively analyzed and optimized in advance. In this scheme, language implementation would consist of two phases. The offline translator generation phase would digest annotations describing the semantics of library routines and combine them with its own analysis to generate an optimized version of the library, and produce a language translator that understands library entry points as language primitives. The script compilation phase would invoke the generated compiler to produce an optimized base language program. The generated compiler would (1) propagate variable property information throughout the script, (2) use a high-level “peepphole” optimizer based on library annotations to replace sequences of calls with faster sequences, and (3) select specialized implementations for each library call based on parameter properties at the point of call.

We will use this strategy to construct a high-level application development environment for an application of interest to Los Alamos based on Matlab. This system could be seen as a flexible replacement for the POOMA framework. POOMA is a parallel object-oriented framework developed at Los Alamos to express scientific programs.

Coordination with Campaigns/DSW: With regard to the milestones discussed below, no indication is given as to the ASCI mileposts supported, since the association of certain ITS milestones with ASCI mileposts presents possible classification issues. In all cases, however, ITS research and development activities are driven by ASCI code needs.

Collaboration and Coordination: We plan to work with the Rapid Performance Prototyping of Algorithms effort (4.2.2.1.3), under the leadership of B. W. VanderHeyden to develop strategies for implementing a PSE based on CartaBlanca. We may also work on the development of a specific replacement for the POOMA framework.

Over the long range, we propose to deliver a prototype version of the Telescoping languages framework based on Matlab and a prototype PSE developed with it. To achieve that goal, we will need to work with an application group at Los Alamos who will provide the domain-specific libraries that will form the basis for the PSE. CartaBlanca is an example of the kind of library that will be needed.

University Contracts: Initially, all the work is being carried out at universities. Rice University is the lead, but key contributions in the area of supporting library design will be provided by University of Houston (L. Johnsson) and Tennessee (J. Dongarra). All of these will be handled through the LACSI academic partners contract with Rice.

Issues, Constraints, and Assumptions: The funding levels for this effort need to be maintained at the current level or grow if development is to accelerate. We will also need to have a high level of interaction with Los Alamos staff. If the SFITL were to be funded, this would be an ideal venue for these interactions.

Milestones and Associated Tasks:

Tasks:

- Develop Matlab front end.
- Develop type analysis system for Matlab system
- Integrate Matlab front end into Rice dHPF analysis and transformation framework.

- Develop language generation driver based on dHPF analysis and transformation system.
- Choose application system for experimentation.
- Adapt domain-specific library for chosen application area to the framework.

- Demonstrate the framework by producing preliminary PSE based on the chosen application.

4.2.2.2.3 Numerical Methods

Sub-Project Lead:

**Danny C. Sorensen, Rice University, Dept. Computational and Applied Mathematics,
(713) 348-5193, sorensen@rice.edu**

Sub-Project Description: The goal of this project is to develop, analyze, and implement parallel computational methods relevant to DOE/Los Alamos applications. Large scale problems are the focus of every activity in this sub-project. The areas being addressed include nonlinear optimization algorithms and software, analysis and optimization of linked subsystems, Krylov projection methods for linear systems and eigenvalue calculations, reduced basis methods for dynamical systems, discrete optimization, methods for solution of nonsmooth, multiscale coupled models, methods and software for inverse and control problems.

Our work is relevant to building computational infrastructure for emerging applications that require coupling of subsystem simulations, and in predictability studies. It will be particularly relevant to predictability of low-dimensional physical events within systems with continuum multifractal structure. Our projects will enable efficient Newton-Krylov methods and stability analysis of complex weather models. The research will support the Los Alamos effort in simulation, identification, and control of complex multiscale phenomena. In many cases we expect to produce preliminary prototypes of object oriented codes relevant Directed Stockpile Work (DSW).

In the area of Linear Solvers (Michele Benzi, Emory University; Miroslav Tuma, CAS, Czech Republic) we hope to develop robust parallel sparse approximate inverse (PAINV) codes. While we currently have a robust implementation for SPD matrices that is well suited for 2D problems, 3-D problems, require a different partitioning strategy (based on Nested Dissection) that needs to be implemented and tested. A nonsymmetric version of the code, which may not be immediately needed by the code teams, is at a more rudimentary stage of development. Additional work will include the development of block versions of PAINV, targeted at systems of PDE's. Part of the research needed is to determine whether this block code would benefit any of the codes of ASCI interest.

Collaboration and Coordination: Kloucek's work at Rice will provide expertise in materials modeling to enhance the physics portfolio addressed by the CartaBlanca program. Sorensen at Rice will work with Nadiga at Los Alamos to develop and implement eigenvalue methods for linear stability analysis of ocean models. Symes will develop techniques for simulation, identification and control of complex multiscale phenomena arising in Los Alamos applications. Optimization, in particular optimization of linked subsystems associated with predictability studies will be a major theme.

University Contracts: Most research work associated with this sub-project is being carried out at Rice University.

Issues, Constraints, and Assumptions: The funding levels for this effort need to be maintained at the current level or grow if development is to accelerate. We will also need to have a high level of interaction with Los Alamos staff. If the SFITL were to be funded, this would be an ideal venue for these interactions.

Milestones and Associated Tasks:

Tasks:

- Incorporation of mesoscale numerical module into Carta Blanca JRPE in addition to the multiphase flow package or examination and modeling of thermodynamics of structured materials.
- Extend subsystem optimization and analysis tools to integrate domain decomposition formulations for single simulations into full system analysis and optimization.
- Develop an eigenvalue analysis technique suitable for performing linear stability analysis on complex ocean models.
- Develop prototypical high performance, reusable, extensible software for inverse and control problems using modern object oriented programming methods.

4.2.2.3 Application Reliability

Project Lead:

Richard Graham, LANL/CCS-1, (505) 665-5685, rlgraham@lanl.gov

Project Overview: This project addresses the ability of ASCI applications to generate reliable solutions. Issues affecting a reliable solution occur through out the solution production system (e.g. in the hardware, software subsystems and in the numerics).

With the increase in the sheer size and complexity of computer hardware and software, issues of application reliability start to become prohibitive problems that must be addressed. These issues include both reliability of the hardware and the software environment and the need to recover from failure in these low level systems. Application numerical accuracy in the context of a very large number of floating point operations is also an issue. This project addresses these types of problems by working on cluster manageability, error anticipation, and the ability to move an application process from hardware that is near failure to other hardware with out interrupting the application. We will look at improving the performance and reliability of the interconnect used for both the LAN and the WAN, as well as at automatic tools for analyzing an application's numerical stability.

4.2.2.3.1 Science Appliance

Sub-Project Lead:

Ronald Minnich, LANL/CCS-1, (505) 667-7239, rminnich@lanl.gov

Sub-Project Description: Work this year will focus on getting the Science Appliance capabilities working on, first, the 128-node Linux Alpha testbed, and, second, the FS-P1 512-node testbed. For the 128-node testbed, the work will include replacing the standard BIOS (the SRM) with the Los Alamos-developed LinuxBIOS. The LinuxBIOS reduces cluster boot time from tens of minutes (standard AlphaServer SC software) to less than 1 minute. The work will include the LinuxBIOS work; extending our work with Scyld Computer Corporation so that their cluster software will function on large-scale clusters; and applications work to ensure that applications run and scale well in a Linux environment.

Coordination with Campaigns/DSW: With regard to the milestones discussed below, no indication is given as to the ASCI mileposts supported, since the association of certain ITS milestones with ASCI mileposts presents possible classification issues. In all cases, however, ITS research and development activities are driven by ASCI code needs.

Collaboration and Coordination: Leveraging DOE Office of Science funds this work encompasses collaborations with IBM, Compaq, Dell, Linuxnetworkx, API networks, Silicon Integrated Systems, VIA, Acer, Linux Laboratories, U. Maryland, and others.

University Contracts: This sub-project funds the development of a scalable cluster debugger at U. Pennsylvania for a debugging environment capable of managing 16,000 or more target processes and the huge amount of data they can produce.

Issues, Constraints, and Assumptions: We're going to need unfettered access to the P1 machine at some point. It is essential that we have around-the-clock access to the large cluster for testing purposes.

Milestones and Associated Tasks:

Tasks:

- Improved Scyld software running on the Linux testbed (FY02, Quarter 1).
 - ZPL working on the Linux testbed (FY02, Quarter 1).
 - Preliminary Parallel Debugging capability on the Linux testbed (FY02, Quarter 2).
 - Demonstration of prototype Linux environment on the FS P1 machine (FY02, Quarter 3).
 - Early "Run Through" testing (FY02, Quarter 4).
-

4.2.2.3.2 Adaptable Reliable SAN and WAN

Sub-Project Leads:

Richard Graham, LANL/CCS-1, (505) 665-5685, rlgraham@lanl.gov

Wu Feng, LANL/CCS-1, (505) 665-2730, feng@lanl.gov

Sub-Project Description: Correct and efficient use of the interconnect by the application is vital for efficient and reliable application performance. Two areas will be addressed in this work, the reliability and adaptability of a message passing library used by end user applications in a SAN environment, and the reliable and efficient transport of user data for storage analysis over the WAN. One of the objectives of the User Level Messaging (ULM) project is to provide a high performance, reliable, adaptable, and scalable message passing environment (MPI being one of it's instantiations). In the context of this work we will be working on adding infrastructure that will allow the messaging layer to support a dynamic NIC environment (such as NIC failure) with out interrupting the application, provided the hardware and the OS don't fail under these conditions. We will also put hooks into the library to support restarting messaging with out killing the application when the OS migrates the application (see section 4.2.2.3.1 above). We will provide an implementation on the Q machine.

With the need to transport large volumes of data, e.g. ASCI visualization data sets, for high-performance computing, networking performance over the wide-area network has become a critical component in such an infrastructure. Unfortunately, because operating systems are still tuned for yesterday's WAN speeds and applications, ASCI visualization researchers at Los Alamos actually find it faster to dump the data sets to tape and use FedEx than to use the network to transfer the information to Sandia! To address this problem, we propose two different and orthogonal solutions that will increase throughput by at least an order of magnitude to the end-user application: [1] dynamic adaptation of flow-control windows that scale beyond gigabit and terabit speeds and [2] a new congestion-control mechanism that scales to today's and tomorrow's networks.

Collaboration and Coordination: We have developed a fault tolerant message passing library (MPI) for the 3T machine, and are in the first stages of getting Los Alamos's ASCI applications up and running with it. A port to the 30T is under way, and we expect this to be the route to provide MPI reliability on this machine. We have started talking with Compaq about how to provide the reliability layer to their Alaska version of MPI.

We have been collaborating with Ken Kennedy, Richard Baraniuk, Edward Knightly, Robert Nowak, and Rudolf Riedi (Rice) via LACSI and Thomas Ndousse (DOE High-Performance Networking) via DOE SciDAC provided the initial impetus of a collaborative effort in network traffic characterization in an effort to better understand the long-range behavior of network traffic so as to provide insight into how to better design high-performance network protocols for the WAN.

Issues, Constraints, and Assumptions: The main issue in the SAN work is the availability of correctly working ELAN3 libraries from Compaq with multiNIC support and the ability of Tru64 to keep the OS up and running when a NIC fails. We are also assuming the availability of

test systems in which we can induce failures, so that we can test the library, as well as timely availability of the 30T machine.

For the WAN work the biggest issue and constraint is Los Alamos's inability to provide an easily accessible, but still secure, open collaboration network (or research network) to test our high-performance networking software. Consequently, we have been jumping through hoops to rig an internal environment to test our software while also talking with SDSC and ANL to have them test our software for us.

Milestones and Associated Tasks:

1. Deliver first version of the adaptable and reliable MPI-1 for the 30T|FS P1. (Q4 FY02)

Tasks:

- Work with ASCI code teams on porting their applications to use the fault tolerant MPI on the P1 system. (FY02, Quarter 1).
- Implement first version of the adaptable and reliable MPI-1 for the 30T. (FY02, Quarter 4)
- Complete version 0.9 of our "dynamic right-sizing" software in the Linux 2.2.x kernel. (FY02, Quarter 2)
- Investigate how experimental, proprietary switching technology could be integrated into a clustered system, which exhibits performance, design and cost characteristics relevant to ASCI's architectural objectives.
- Complete initial investigation on new congestion-control mechanisms (FY02, Quarter 3)
- Complete version 1.0 of our "dynamic right-sizing" software in the Linux 2.4.x kernel. (FY02, Quarter 4)

4.2.2.3.3 Code-based Sensitivity Analysis

Project Lead:

Mike Fagan, Rice University, (713) 348-5178, mfagan@cs.rice.edu

Project Overview: Users of complex, large-scale computer models often need sensitivity information in addition to model outputs. Sensitivity calculations are at the crux of many uncertainty quantification methods. Other important calculations that use sensitivity information include: round off error estimation, parameter identification, model fitting, data assimilation and data fusion, generalized branch-and-bound, and optimal design.

In general, there are two different ways to compute desired sensitivities:

1. Evaluate the model over a “spread” of input values, and compute the sensitivities from the spread of model output values.
2. Develop a new program that computes the sensitivities of the original model.

The “spread” method is often inaccurate, and exhibits poor performance. The “new program” method is (should be) more accurate and has (should have) good performance, but the development of such a program takes a prohibitive amount time, if done by human developers. *Code-based sensitivity analysis* is a set of techniques for automating the “new program” method. By combining modern compiler-based source code analysis and transformation with modern numerical analysis algorithms, the Code-based sensitivity analysis project strives to develop tools that automatically transform computer models expressed in high-level source code to *augmented* computer models that compute designated sensitivities in addition to model outputs. Code-based sensitivity augmentation produces accurate, good performing sensitivity-enhanced codes in much the same way that modern optimizing compilers produce good assembly language programs.

This project is based on our previous success with the Adifor 3.0 automatic differentiation project. We plan to generalize the Adifor system in two different aspects:

1. Computer language: The Adifor 3.0 system accepts only Fortran 77. We wish to extend the system to Fortran 90, C, C++ and Java.
2. Sensitivity calculations: Adifor 3.0 currently supports gradients, adjoints and some limited Hessian capability. We wish to extend the sensitivity calculation to arbitrary higher order derivatives, intervals, probability distributions, and some novel sensitivity notions based on Fourier coefficients or wavelets.

Issues, Constraints, and Assumptions:

- Need to identify the “F90 short list” of 3-4 F90 features, none of which are very different from current vendor-supported Fortran 77 extensions.

- The acceptance tests do *not* have to be security-sensitive ASCII codes. People who know these codes, however, should definitely have input to ensure that the acceptance tests model the features of interest to the ASCII project.
- Meeting the Q4 '02 deadline is based on the assumption that refactoring the current Adifor augmentation engine does not take significant time (current estimate is 2.5 - 3 man months for this task).

Milestones and Associated Tasks:

Tasks:

- Agree with Los Alamos upon an Adifor "F90 short list" of features.(Q1 '02)
- Design with Los Alamos an automatable acceptance test suite for the Adifor "F90 short list" sensitivity tool. (Q2 '02)
- Refactor current Adifor sensitivity tool to accommodate new augmentation infrastructure. (FY02, Quarter 4)
- Add "F90 short list" features to refactored Adifor to produce a prototype "F90 short list" sensitivity tool. (FY02, Quarter 4)

4.2.2.4 Computer Science Community Interaction

Project Lead:

Andrew White, LANL/CCS-DO, (505) 665-4700, abw@lanl.gov

Project Overview: For LACSI to be successful, it must build strong academic ties to experts in the community able to further the strategic goals of the Institute in support of the ASCII program. Our collaborations include Rice University, University of Houston, University of Tennessee, University of New Mexico and University of Illinois–Urbana Champaign. The past several years have seen several highly successful projects in this area including; activities which brought world-class computer science researchers and graduate fellows to Los Alamos, interactions that presented ASCII work in conjunction with well respected universities, and research exposure to the community at the Annual LACSI Symposium. The present year's sub-projects, described below, continue to build ASCII's interaction with the computer science community.

4.2.2.4.1 Rice University's Santa Fe Information Technology Laboratory (SFITL)**Sub-Project Leads:**

Andy White, LANL/CCS-DO, (505) 665-4700, abw@lanl.gov

Ken Kennedy, Rice University, Center for High Performance Software Research,
(713) 348-5186, ken@rice.edu

Sub-Project Description: The Santa Fe Information Technology Laboratory (SFITL) is a world-class computer science research facility to be established in Santa Fe, New Mexico. It would leverage the campus-based research efforts currently supported by the DOE ASCI program to provide a unique focal point for nonclassified research collaborations between the DOE national laboratories, especially those involved in the Advanced Strategic Computing Initiative (ASCI) program (Los Alamos, Livermore, and Sandia), and ASCI academic partners.

This project is intended to create the world's premiere academic laboratory for research on software for high-performance technical computing, and to convey the results of that research in the form of knowledge and software artifacts to the DOE laboratories and the national HPCC community in support of DOE activities. Activities at the SFITL would include:

- *programmatic research* on a variety of topics in high-performance computation;
- *opportunity-driven research* that would address important new research issues in high-performance technical computing as they arise;
- *educational outreach* to the national laboratories and to the national HPCC community; and
- *collaborative activities*, including joint projects with industry, visiting faculty programs, and undergraduate and graduate research.

Rice University has developed a long-term business plan for the SFITL. The key components of this plan are sustained core DOE support for at least five years and Rice's plans to construct a permanent facility in Santa Fe, currently planned for the fourth year of the contract. As a part of this proposal, Rice University commits to keeping SFITL operational so long as DOE funding remains at the proposed levels. Beyond that, Rice will maintain the laboratory for as long it remains economically feasible.

In addition to carrying out the research and education programs, Rice will take strong steps to secure *academic partners*—institutions with strong research programs in high-performance computing—to participate in SFITL activities by sending visiting faculty and staff. In addition it will seek active industrial participation in the lab, particularly from those companies with substantial interests in computing within Department of Energy programs.

Los Alamos will take delivery of the most powerful, and complex, computer ever assembled in early 2002. The challenge of making this, or any other modern high-performance computer, an effective scientific and engineering tool are enormous; the size of this challenge is exceeded only by the size of the reward if we are successful. The SFITL will focus the external academic community on the computer science issues associated with computing at this scale: for example,

performance, software environment, communications, parallel algorithms, and reliability. Computer and computational science are clearly critical skills for Stockpile Stewardship.

The goals of the SFITL are summarized as follows.

- establish a center of excellence in computer science in Northern New Mexico that contributes to DOE mission goals;
- pursue computer science research that is relevant to the goals of High-Performance Computing (HPC) programs of DOE and the national laboratories;
- establish a unique laboratory focused on advanced research and professional and graduate education in computer science and information technology;
- provide a strong focus on high-performance computing in the academic computer science community;
- provide opportunities for students and faculty, especially from academic partners of the ASCI laboratories, to participate in high-tech research and educational collaborations;
- provide new professional and graduate educational opportunities for Northern New Mexico in computer science and information technology;
- provide a locus for national and international collaborations in computer science and information technology;
- foster knowledge and technology transfer among enterprises in Northern NM, thus supporting the emergence of Santa Fe as an information technology center; and
- provide a mechanism for collaborations with industry.

Structure: SFITL will be an extension of Rice University reporting to the Office of the Provost. SFITL will be strategically located in Santa Fe, New Mexico, roughly halfway between Los Alamos National Laboratory and Sandia National Laboratory in Albuquerque. This is close enough to both laboratories and the University of New Mexico in Albuquerque to foster close collaborative relationships. Summer Workshops and long-term visitor programs will encourage collaboration with Livermore National Laboratory, the many academic partner institutions involved in the DOE ASCI research program, and other collaborators located outside of New Mexico. In addition, the SFITL will forge strong ties to the emerging information technology community in Northern New Mexico. Initially, the programs supported by the SFITL will be focused on topics in high-performance computing that leverage the strengths of Rice and other academic institutions involved in the DOE ASCI program. As the SFITL matures, and stabilizes financially, the scope of the programs will increase to include a broader range of subjects within information technology and related sciences.

Functionally, SFITL will be an extension of the Houston campus of Rice University. Rice will own and operate the facilities, and all staff paid through the SFITL will be employees of Rice. Rice will solicit *academic partners* for the endeavor that will add value to the research enterprise by participating in SFITL research through its visiting program. Most of these academic partners will be institutions currently involved in ASCI-supported HPC research. Finally, Los Alamos, Sandia, and Livermore national laboratories will be able to interact with SFITL much as an

academic partner would. In particular, these laboratories will be able to designate a number of their own staff members to spend time at SFITL working on joint research. This arrangement will involve an additional overhead fee to cover the cost of space for long-term visitors

To reduce the impact of the distances among the partners, SFITL will be committed to using advanced facilities, *e.g.* an Access Grid node, for remote collaboration and education. Specific technological choices will be compatible with teleconferencing and telecollaboration facilities at the national laboratories.

SFITL will also establish a program of industrial partnerships that will involve companies in the research carried out under this proposal and other proposals that may be forthcoming. Because of its importance to the construction of the 30-teraops supercomputer at Los Alamos, a particularly high priority will initially be given to attracting a research investment from Compaq.

To achieve its goals, SFITL must be a highly desirable place to visit and work. To this end, it must provide a high-quality intellectual environment and have an appealing location. For the first two years, Rice University will rent space for SFITL near Santa Fe's historic downtown district. Under the current plan, SFITL will move into a new building designed for it and constructed on Rice property in Santa Fe in the fourth year. Throughout its lifetime, the SFITL will have a high quality computing and communications environment and access to additional resources at its partner institutions.

The activities of the SFITL will be organized into three major categories: (1) Project-Oriented Research and Development, (2) Opportunity-Driven Research, and (3) Outreach, Education, and Collaboration. These are not intended to be orthogonal or even disjoint. It is expected that most of the professional staff and visitors to the ITL will be involved in more than one of these areas.

Project-Oriented Research and Development: The proposed funding will support multiperson, multiinstitution research efforts in computer science. It will be uniquely situated to foster collaborations among industry, academia, and DOE's laboratories. Under SFITL, funded projects will involve long-term fundamental research on enabling technologies and applications of high-performance computing. The strategy will be to carry out research that is focused on long-term goals, but can produce useful research artifacts—software tools, libraries, and prototype applications—that can be useful to the DOE program in the near and medium term.

Initially, research would concentrate on the major themes of the Los Alamos Computer Science Institute, with which Rice University is already affiliated, but which are highly relevant to all of DOE's high-end computing needs:

1. Software for *future-generation computing platforms*, including PC clusters and ultrascale systems (*e.g.*, processor-in-memory architectures, Blue Gene). This work would address the general and architecture-specific challenges of scaling to a petaflop and beyond, including managing node-level memory hierarchies, and communication minimization. A major byproduct of this work would be software that would provide support for the ASCI 30 and 100 teraops systems.

2. Software support for *retargetable high-performance applications*, including self-tuning applications, cache-oblivious algorithms, and portable run-time systems. This effort would seek to use careful structuring and high levels of computational power to retarget applications to new scalable architectures without requiring enormous investments of programmer time.
3. *High-level programming systems*, including component frameworks for generation of problem-solving environments, advanced data structure libraries, and domain-specific libraries. A principal goal of this work is to produce frameworks that help generate high-level domain-specific programming languages from annotated component libraries. The goal is to achieve performance that is competitive with hand-coding the application from scratch.
4. *Computational mathematics*, including algorithms and mathematical software for scalable parallel and distributed heterogeneous systems. Areas of interest would include CFD, domain decomposition, sparse and dense linear algebra, linear and nonlinear optimization, fast Fourier transforms, and integer and mixed integer programming. A major goal of this effort would be to produce libraries that could support the other research efforts, especially retargetable applications and high-level programming systems.
5. Software to support application development for *computational Grids*. This work would build on systems being developed for the GrADS (Grid Application Development Software) project by expanding programming systems to support applications of interest to DOE and the national laboratories.

In addition to these specific areas, the SFITL will actively collaborate in community-based open-source software development activities and will actively seek to release its own software for use in such efforts, where this software is not covered by other licensing agreements.

Opportunity-Driven Research: To keep SFITL at the forefront of computer science research it is vital that it provide an intellectual environment that will attract the best people and ensure that those people generate a continual flow of exciting new ideas. To achieve this goal, SFITL must:

- provide an attractive and exciting destination that will attract world-class visitors;
- provide a stimulating environment for resident staff, DOE laboratory staff, post-docs, and visitors;
- support “skunk works” research by individuals and small groups, with funding for targets of opportunity;
- provide an attractive site for brainstorming sessions and workshops that identify critical problems and approaches to solving them; and
- provide staff support for meetings, lecture series, and conferences of importance to SFITL and DOE.

To these ends, SFITL will set aside funding for *opportunity-driven research* that explores new directions as they arise. The results of these efforts will be ideas, publications, and proposals. In addition, some efforts will evolve into project-oriented research. Opportunity-driven activities

must be consistent with the overall mission of SFITL. A small number of “organizing themes” can be used to achieve this. These themes could be established each year as opportunities arise. For example, two themes that might be used in the early years are:

- *Ultrascale algorithmics*. This theme would cover all aspects of building and using effective ultrascale computer systems. In the near term, the Q machine and related problems will provide a focus.
- *Critical simulations*. This theme involves the application of scientific/engineering simulations to socially important and time critical problems such crisis management.

Outreach, Education, and Collaboration: To meet the goal of providing an intellectual focus for information technology in Northern New Mexico, SFITL will conduct several programs to foster education, technical communication, and technology transfer. While some aspects will be to serve the needs of DOE and the national laboratories, these programs will benefit the nation, New Mexico, and Santa Fe.

- In the first years of operation, we will establish a *SFITL Short Course* series that will bring faculty from Rice and its academic partners to Santa Fe to offer short courses (up to one week) on topics of interest to the DOE national laboratories. Such topics would range from compilers and operating systems for scalable systems to algorithms for specific types of problems in computational science.
- In the first year, SFITL will establish a *Distinguished Lecture Series* that would bring outstanding computer and computational scientists to conduct seminars on their topic of specialization. Some of these lectures would be given at SFITL and some would be offered at Los Alamos and Sandia. As teleconferencing facilities become operational, the lectures would be transmitted to academic partners and other national laboratories.
- In the first full year of operation, we will establish a program of *Summer Internships in High-Performance Computing* for graduate and undergraduate students at SFITL academic partners, designed to provide the student with experience working on research in high-performance computation. These internships would be open to all students and would be awarded based on an application that includes a research prospectus.
- By the second year of operation, we will establish two *SFITL Summer Seminar* programs. Patterned after the extremely successful NATO Summer School programs, these seminars will include a *Research Seminar* series, intended to advance the state of the art in important topics, and a *Study Seminar* series, intended to bring researchers up to speed on those topics. Nationally known scientists will be recruited to organize the technical programs while SFITL will provide administrative and logistical support.
- Eventually, SFITL would consider expanding its educational programs to include graduate courses in computer science and computational science from Rice and SFITL academic partners.

In addition to these efforts, the SFITL will initiate outreach efforts designed to increase the participation of underrepresented groups in science and engineering, following the successful model pioneered by the Center for Research on Parallel Computation. SFITL programs would include summer research programs for students from underrepresented groups and technology

education programs for New Mexico schoolteachers. These programs would incorporate discussion of how to increase interest in information technologies from minorities and women.

Management and Planning: Management of the SFITL would be the responsibility of the *Director* (Ken Kennedy) who reports to the Provost of Rice University. In carrying out this responsibility, the Director will be assisted by the *Chief Scientist*, who will be responsible for planning and who will serve as Deputy Director of SFITL. Either the Director or the Deputy Director will be in residence at the SFITL. The day-to-day administration will be handled by the *Executive Director* (Rob Fowler), who will be full-time at the Santa Fe location, and the *Associate Director for Business and Administration* (Danny Powell). Both the Executive Director and the Associate Director for Business will report to the Director. There will be an *Executive Committee* to serve as the principal planning body, advising the Director on short- and long-term strategic plans. In addition to the four administrative positions named above, the Executive Committee will include representatives from each stakeholder with staff located at the SFITL and from each major research thrust group. An *Oversight Committee* will additionally have representatives from each of the institutional stakeholders (e.g., Rice and participating ASCI national laboratories) and from major academic partner institutions. Finally, SFITL will establish an *External Advisory Committee* to provide the director with guidance on the development of research, education and outreach programs. This committee will include a distinguished and diverse group of representatives from industry, academia, and government.

Issues, Constraints, and Assumptions: Assumptions are that the core DOE funding level will be established at \$5 million per year and that due consideration will be given to continued funding beyond the five-year minimum lifetime. Further assumptions are that there be a two-year phase-out period if a decision is made to discontinue funding at any time. This would permit Rice the opportunity to shut down the laboratory if other funding is not secured. In addition, the on-campus ASCI funding for Rice and other academic partners need to be maintained throughout the lifetime of the contract. This is essential to maintain a high level of faculty attention from academic partners. Finally, we are assuming that DOE will support a significant number of staff members from Los Alamos, Sandia, Livermore, and other national laboratories to conduct research on-site at SFITL. SFITL would provide space for these researchers at modest overhead charges to the originating laboratories, which would be in addition to the proposed research funding.

Milestones and Associated Tasks:

Tasks:

- Create and implement a staffing and benefits plan.
- Complete appropriate contractual arrangements with DOE.
- Locate, negotiate, and obtain initial facilities—including space, office furnishings, networking, and computing equipment.

- Advertise, recruit, and hire initial staff

- Initiate industrial partners program.
- Initiate visiting scientist and Distinguished Lecture programs targeting DOE, academic, and industrial researchers.
- Initiate research programs.

4.2.2.4.2 High-Performance Computer Science Fellowships

Sub-Project Lead:

Rod Oldehoeft, LANL/CCS-1, (505) 665-3663, rro@lanl.gov

Sub-Project Description: The High-Performance Computer Science (HPCS) Fellowship Program supports Ph.D.-level computer science students pursuing research in areas relevant to the needs of ASCI. The Program is modeled after the highly successful DOE Computational Science Graduate Fellowships. In addition to pursuing graduate research, each supported fellow will do a practicum (generally during the summer) at Los Alamos working with LACSI/Los Alamos researchers. LACSI/Los Alamos researchers may participate as external members of student advisory committees. In FY02 two additional HPCS fellows will begin support, bringing our total to six fellows.

Collaboration and Coordination: The HPCS Fellowship Program is a tri-lab collaboration, with additional fellows funded by the Computer Science Institutes at Lawrence Livermore and Sandia. The overall Program is administered by Krell Institute, which has contracts with each laboratory.

Issues, Constraints, and Assumptions: The Program is designed to provide up to four years of support (renewable annually based on satisfactory progress) for each fellow, so long-term planning is required to provide the necessary continuity.

Milestones and Associated Tasks:

Tasks:

- Develop a pool of qualified applicants for new fellowships.
- Evaluate applicants, select fundable potential fellows.
- Evaluate continuing fellows for ongoing support.
- Manage practica for fellows visiting Los Alamos.

4.2.2.4.3 Collaboration and Management

Sub-Project Leads:

**John Thorp, LANL/CCS-1, (505) 665-8226, thorp@lanl.gov
Danny Powell, Rice University, (713) 348-6011, danny@rice.edu**

Sub-Project Description: LACSI is a collaborative research effort between Los Alamos, Rice University, University of Illinois at Urbana-Champaign, University of Houston, and the University of Tennessee at Knoxville. Effective means of supporting collaborations are important to the success of the Institute. To do this, this project will support a variety of opportunities for researchers from Los Alamos and the academic partner sites to visit each other, to share ideas and to actively collaborate on technical projects. In addition, we will organize, host, and otherwise support a series of technical workshops as well as an Annual LACSI Symposium, not only to showcase LACSI results, but also to provide a forum for presenting outstanding research results from the national community in areas overlapping the LACSI technical vision. We will also coordinate a technical infrastructure between Los Alamos and the academic partners, enabling web broadcasting of local technical talks, workshops, and the Annual Symposium to an off-site audience.

Collaboration and Coordination: The work will require coordination with the LACSI researchers at the partner sites and with their administrative and IT staffs.

Milestones and Associated Tasks:

1. Complete formal three year review of the Rice University (et. al.) research contract —FY02, Quarter 1.

Tasks:

- Coordinate a common technical infrastructure at Los Alamos and the academic sites to support web collaboration of relevant research talks and workshops.
- Organize, host, and/or support high quality technical workshops of interest to the LACSI mission and to ASCI.
- Host and support a LACSI Speaker Series at Los Alamos (hosting academic speakers) and at the academic sites (hosting Los Alamos speakers).
- Support summer and academic year collaborative research project visits between Los Alamos and academic researchers.
- Manage the operation of the Institute.
- Organize and host a high quality, Annual LACSI Symposium, with distributable Proceedings of interest to a broad computational science research community.